

Maritime notes

21. april 2021 16:21

Vi er meget interesseret i at få nogle KPI på hvor godt vi rammer med vores ETA beregning.

Loggen indeholder en strib records der bliver dumpet hver 15 minutter og når der sker ændringer.

Kolonnerne er:

track_id, mmsi, status, port_id, shape_id, stamp, eta_erp, eta_ais, ata_ais, bs_ts, sog, username

track_id knytter turen sammen.

mmsi er skibets nummerplade

status er en "Tracking code" der kan ses i bunden af den her: <https://maritime.gatehouse.dk/pub/api-doc/portserver/> hvor 14 er ankommet

port_id / shape_id er en id på den destination (havn eller location) hvor skibet sejler hen.

stamp er tidspunktet for log entry

eta_erp er det som skibet selv har oplyst som forventet ankomst tid

eta_ais er vores beregnede ankomst tid (den vi gerne vil have checket)

ata_ais er den faktisk ankomst tid når status skifter til 14 (ankommet)

bs_ts er timestamp for sidste AIS besked (nok ikke så vigtig)

sog er Speed over ground på tidspunktet (bs_ts)

username er brugernavnet.

```
# Vi vil i sidste ende gerne kunne regne mean(|eta - ata|)  
# parametre vi vil teste  
# eta1  
# givet et skib og hvor det er på vej hen  
# det beregner en rute og givet nogle historiske hastigheder beregnes en eta.  
# Givet en log hvor godt rammer vi tidspunktet  
# 90 % bedste bud hvor godt rammer vi ata,  
# Hvor godt performer vi for erp_eta.  
# eta2  
# Containerskibe - kan lave eta på en hele rute  
# destination predictor
```

Tracking Codes

The tracking code is an identifier assigned for each ship being tracked.

Code	Message	Description
0		Idle or initialized. The track has not yet been picked up. Will typically happen right after initiating the track and if the system is under heavy load.
1		Tracked. This is the normal state for a ship being tracked towards a destination.
2		Predicted. The ship has been picked up by the tracker, but have not had valid position updates for > 15 minutes. This is typical in areas covered by satellite.
3		Too slow. The ship is being tracked but is laying still. It does not make sense to attempt to calculate an ETA.
9		End of predicted route. The ship has been predicted to have reached its destination, but no valid position updates > 15 minutes.
12	Arriving	The ship is arriving at the destination. For a port that means it is entering the port area but is still at speed.
14	Arrived	The ship has arrived at the destination. For a port that means it is laying still inside the port. The tracking is stopped
16	Departed	The ship has departed the destination. The tracking is stopped

18	Canceled	The ship track has been canceled
19	Expired	No additional AIS reports have been received for too long. The ship track is considered lost
21		Source not found
22		Destination not found
23		Route not found. Source and destination were found, but it was not possible to find a route. Could be due to draught
24		Invalid Track
29		System Error

Fra <<https://maritime.gatehouse.dk/pub/api-doc/portserver/>>

Data cleaning

Vi er interesseret i at sammenligne eta_ais med ata_ais. Derfor skal vi have fjernet alt det data for skibe der enten ikke er ankommet eller som i perioden kun er logget som ankommet.

Derudover logges ruten flere gange efter skibet er ankommet. Der er eksempler på at eta_erp opdateres efter skibet er ankommet. Vi fjerner dog alle logs med status 14 på nær den første.

Suez kanalen - blokade

23. marts til 30 marts iflg. <https://www.dk.dsv.com/om-dsv/Presse/News/2021/Marts/blokaden-af-Suezkanalen-vil-paavirke-spefragtsmarkedet-i-maaneder-frem>

Ideer:

- Man kunne lave et plot af prædiktionen vs tidspunkt prædiktionen er fundet sted. Det ville nok illustrativt vise performance.
- Man kan lave et plot over den gennemsnitlige absolute fejl vs tidspunkt.
- Vi kan estimere grand mean af absolute error.
- Undersøge over og under estimering af eta

Resultater

Grand mean and standard deviation

	MAE ETA_ais	Std ETA_ais	MAE ETA_erp	Std ETA_erp
all	37.89	56.33	6.05	21.66
Before 23/03	37.80	54.39	6.23	24.19
After 23/03	17.48	25.12	3.79	8.38

Tallene i de to nederste rækker er regnet ved at betragte data med timestamps i de relevante perioder. **Man burde nok kigge på ata i stedet for stamp**.

Det ses at ETA er bedre efter 23/03 hvilket kan skyldes at kortere ture er betragtet. Disse resultater kan nok ikke bruges til så meget som de står nu.

I stedet for at se på alt data, vil vi nu se på de 90% bedste ETA_ais:

Hvordan definerer vi de 90% bedste?

1. Regner absolut fejl for alle time stamps og tager de 90% mindste.

2. Regner absolut fejl for alle time stamps og vælger 90% af track_ids med mindste fejl.

Ad 1. får vi (bemærk vi har brugt alt data for ETA_erp og ikke kun 90% bedste)

%	MAE ETA_ais	Std ETA_ais	MAE ETA_erp	Std ETA_erp
90%	22.47	28.06	6.055	21.66
50%	3.24	3.32		